

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 09/753,008A  
Source: IFW16  
Date Processed by STIC: 12/14/2005

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 12/14/2005

PATENT APPLICATION: US/09/753,008A

TIME: 12:15:52

Input Set : A:\96700-658.txt

Output Set: N:\CRF4\12142005\I753008A.raw

3 <110> APPLICANT: Somlo, Stefan  
 4 Mochizuki, Toshio  
 6 <120> TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE PKD2 GENE AND USES  
 THEREOF  
 8 <130> FILE REFERENCE: 96700/658  
 10 <140> CURRENT APPLICATION NUMBER: 09/753,008A  
 11 <141> CURRENT FILING DATE: 2001-01-02  
 13 <150> PRIOR APPLICATION NUMBER: 08/651,999  
 14 <151> PRIOR FILING DATE: 1996-05-23  
 16 <150> PRIOR APPLICATION NUMBER: 09/385,752  
 17 <151> PRIOR FILING DATE: 1999-08-30  
 19 <160> NUMBER OF SEQ ID NOS: 15  
 21 <170> SOFTWARE: PatentIn version 3.3  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 866  
 25 <212> TYPE: PRT  
 26 <213> ORGANISM: Homo sapiens  
 29 <220> FEATURE:  
 30 <221> NAME/KEY: misc\_feature  
 31 <222> LOCATION: (384)..(384)  
 32 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 34 <220> FEATURE:  
 35 <221> NAME/KEY: misc\_feature  
 36 <222> LOCATION: (395)..(395)  
 37 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 39 <220> FEATURE:  
 40 <221> NAME/KEY: misc\_feature  
 41 <222> LOCATION: (426)..(426)  
 42 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 44 <220> FEATURE:  
 45 <221> NAME/KEY: misc\_feature  
 46 <222> LOCATION: (432)..(432)  
 47 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 49 <220> FEATURE:  
 50 <221> NAME/KEY: misc\_feature  
 51 <222> LOCATION: (547)..(547)  
 52 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 54 <220> FEATURE:  
 55 <221> NAME/KEY: misc\_feature  
 56 <222> LOCATION: (586)..(588)  
 57 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 59 <220> FEATURE:  
 60 <221> NAME/KEY: misc\_feature  
 61 <222> LOCATION: (849)..(849)

CP9-6)

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/753,008A

DATE: 12/14/2005

TIME: 12:15:52

Input Set : A:\96700-658.txt

Output Set: N:\CRF4\12142005\I753008A.raw

62 &lt;223&gt; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

64 &lt;400&gt; SEQUENCE: 1

```

66 Met Val Asn Ser Ser Arg Val Gln Pro Gln Gln Pro Gly Asp Ala Lys
67 1      5      10      15
70 Arg Pro Pro Ala Pro Arg Ala Pro Asp Pro Gly Arg Leu Met Ala Gly
71      20      25      30
74 Cys Ala Ala Val Gly Ala Ser Leu Ala Ala Pro Gly Gly Leu Cys Glu
75      35      40      45
78 Gln Arg Gly Leu Glu Ile Glu Met Gln Arg Ile Arg Gln Ala Ala Ala
79      50      55      60
82 Arg Asp Pro Pro Ala Gly Ala Ala Ala Ser Pro Ser Pro Pro Leu Ser
83 65      70      75      80
86 Ser Cys Ser Arg Gln Ala Trp Ser Arg Asp Asn Pro Gly Glu Glu Glu
87      85      90      95
90 Ala Glu Glu Glu Glu Glu Val Glu Gly Glu Glu Gly Gly Met Val
91      100     105     110
94 Val Glu Met Asp Val Glu Trp Arg Pro Gly Ser Arg Arg Ser Ala Ala
95      115     120     125
98 Ser Ser Ala Val Ser Ser Val Gly Ala Arg Ser Arg Gly Leu Gly Gly
99      130     135     140
102 Tyr His Gly Ala Gly His Pro Ser Gly Arg Arg Arg Arg Arg Glu Asp
103 145     150     155     160
106 Gln Gly Pro Pro Cys Pro Ser Pro Val Gly Gly Gly Asp Pro Leu His
107      165     170     175
110 Arg His Leu Pro Leu Glu Gly Gln Pro Pro Arg Val Ala Trp Ala Glu
111      180     185     190
114 Arg Leu Val Arg Gly Leu Arg Gly Leu Trp Gly Thr Arg Leu Met Glu
115      195     200     205
118 Glu Ser Ser Thr Asn Arg Glu Lys Tyr Leu Lys Ser Val Leu Arg Glu
119      210     215     220
122 Leu Val Thr Tyr Leu Leu Phe Leu Ile Val Leu Cys Ile Leu Thr Tyr
123 225     230     235     240
126 Gly Thr Glu Ala Asp Asn Arg Ser Phe Ile Phe Tyr Glu Asn Leu Leu
127      245     250     255
130 Leu Gly Val Pro Arg Ile Arg Gln Leu Arg Val Arg Asn Gly Ser Cys
131      260     265     270
134 Ser Ile Pro Gln Asp Leu Arg Asp Glu Ile Lys Glu Cys Tyr Asp Val
135      275     280     285
138 Tyr Glu Thr Ala Ala Gln Val Ala Ser Leu Lys Lys Asn Val Trp Leu
139      290     295     300
142 Asp Arg Gly Thr Arg Ala Thr Phe Ile Asp Phe Ser Val Tyr Asn Ala
143 305     310     315     320
146 Asn Ile Asn Leu Phe Cys Val Val Arg Leu Leu Val Glu Phe Pro Ala
147      325     330     335
150 Thr Gly Gly Val Ile Pro Ser Trp Gln Phe Gln Pro Leu Lys Leu Ile
151      340     345     350
154 Arg Tyr Val Thr Thr Phe Asp Phe Phe Leu Ala Ala Cys Glu Ile Ile
155      355     360     365
W--> 158 Phe Cys Phe Phe Ile Phe Tyr Tyr Val Val Glu Glu Ile Leu Glu Xaa

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/753,008A

DATE: 12/14/2005

TIME: 12:15:52

Input Set : A:\96700-658.txt

Output Set: N:\CRF4\12142005\I753008A.raw

```

159      370      375      380
W--> 162 Ile Arg Ile His Lys Leu His Tyr Phe Arg Xaa Ser Phe Trp Asn Cys
163 385      390      395      400
166 Leu Asp Val Val Ile Val Val Leu Ser Val Val Ala Ile Gly Ile Asn
167      405      410      415
W--> 170 Ile Tyr Arg Thr Ser Asn Val Glu Val Xaa Leu Leu Gln Phe Leu Xaa
171      420      425      430
174 Glu Asp Gln Asn Thr Phe Pro Asn Phe Glu His Leu Ala Tyr Trp Gln
175      435      440      445
178 Ile Gln Phe Asn Asn Ile Ala Ala Val Thr Val Phe Phe Val Trp Ile
179      450      455      460
182 Lys Leu Phe Lys Phe Ile Asn Phe Asn Arg Thr Met Ser Gln Leu Ser
183 465      470      475      480
186 Thr Thr Met Ser Arg Cys Ala Lys Asp Leu Phe Gly Phe Ala Ile Met
187      485      490      495
190 Phe Phe Ile Ile Phe Leu Ala Tyr Ala Gln Leu Ala Tyr Leu Val Phe
191      500      505      510
194 Gly Thr Gln Val Asp Asp Phe Ser Thr Phe Gln Glu Cys Ile Phe Thr
195      515      520      525
198 Gln Phe Arg Ile Ile Leu Gly Asp Ile Asn Phe Ala Glu Ile Glu Glu
199      530      535      540
W--> 202 Ala Asn Xaa Arg Val Leu Gly Pro Ile Tyr Phe Thr Thr Phe Val Phe
203 545      550      555      560
206 Phe Met Phe Phe Ile Leu Leu Asn Met Phe Leu Ala Ile Ile Asn Asp
207      565      570      575
W--> 210 Thr Tyr Ser Glu Val Lys Ser Asp Leu Xaa Xaa Xaa Ala Gln Gln Lys
211      580      585      590
214 Ala Glu Met Glu Leu Ser Asp Leu Ile Arg Lys Gly Tyr His Lys Ala
215      595      600      605
218 Leu Val Lys Leu Lys Leu Lys Lys Asn Thr Val Asp Asp Ile Ser Glu
219      610      615      620
222 Ser Leu Arg Gln Gly Gly Gly Lys Leu Asn Phe Asp Glu Leu Arg Gln
223 625      630      635      640
226 Asp Leu Lys Gly Lys Gly His Thr Asp Ala Glu Ile Glu Ala Ile Phe
227      645      650      655
230 Thr Lys Tyr Asp Gln Asp Gly Asp Gln Glu Leu Thr Glu His Glu His
231      660      665      670
234 Gln Gln Met Arg Asp Asp Leu Glu Lys Glu Arg Glu Asp Leu Asp Leu
235      675      680      685
238 Asp His Ser Ser Leu Pro Arg Pro Met Ser Ser Arg Ser Phe Pro Arg
239      690      695      700
242 Ser Leu Asp Asp Ser Glu Glu Asp Asp Asp Glu Asp Ser Gly His Ser
243 705      710      715      720
246 Ser Arg Arg Arg Gly Ser Ile Ser Ser Gly Val Ser Tyr Glu Glu Phe
247      725      730      735
250 Gln Val Leu Val Arg Arg Val Asp Arg Met Glu His Ser Ile Gly Ser
251      740      745      750
254 Ile Val Ser Lys Ile Asp Ala Val Ile Val Lys Leu Glu Ile Met Glu
255      755      760      765

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/753,008A

DATE: 12/14/2005

TIME: 12:15:52

Input Set : A:\96700-658.txt

Output Set: N:\CRF4\12142005\I753008A.raw

```

258 Arg Ala Lys Leu Lys Arg Arg Glu Val Leu Gly Arg Leu Leu Asp Gly
259      770                      775                      780
262 Val Ala Glu Asp Glu Arg Leu Gly Arg Asp Ser Glu Ile His Arg Glu
263 785                      790                      795                      800
266 Gln Met Glu Arg Leu Val Arg Glu Glu Leu Glu Arg Trp Glu Ser Asp
267                      805                      810                      815
270 Asp Ala Ala Ser Gln Ile Ser His Gly Leu Gly Thr Pro Val Gly Leu
271                      820                      825                      830
274 Asn Gly Gln Pro Arg Pro Arg Ser Ser Arg Pro Ser Ser Ser Gln Ser
275                      835                      840                      845
W--> 278 Xaa Thr Glu Gly Met Glu Gly Ala Gly Gly Asn Gly Ser Ser Asn Val
279      850                      855                      860
282 His Val
283 865
286 <210> SEQ ID NO: 2
287 <211> LENGTH: 523
288 <212> TYPE: PRT
289 <213> ORGANISM: Homo sapiens
292 <220> FEATURE:
293 <221> NAME/KEY: misc_feature
294 <222> LOCATION: (46)..(46)
295 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
297 <220> FEATURE:
298 <221> NAME/KEY: misc_feature
299 <222> LOCATION: (51)..(55)
300 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
302 <220> FEATURE:
303 <221> NAME/KEY: misc_feature
304 <222> LOCATION: (89)..(93)
305 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
307 <220> FEATURE:
308 <221> NAME/KEY: misc_feature
309 <222> LOCATION: (101)..(111)
310 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
312 <220> FEATURE:
313 <221> NAME/KEY: misc_feature
314 <222> LOCATION: (166)..(167)
315 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
317 <220> FEATURE:
318 <221> NAME/KEY: misc_feature
319 <222> LOCATION: (293)..(293)
320 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
322 <220> FEATURE:
323 <221> NAME/KEY: misc_feature
324 <222> LOCATION: (334)..(335)
325 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
327 <220> FEATURE:
328 <221> NAME/KEY: misc_feature
329 <222> LOCATION: (437)..(440)

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/753,008A

DATE: 12/14/2005

TIME: 12:15:52

Input Set : A:\96700-658.txt

Output Set: N:\CRF4\12142005\I753008A.raw

```

330 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
332 <220> FEATURE:
333 <221> NAME/KEY: misc_feature
334 <222> LOCATION: (511)..(517)
335 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
337 <400> SEQUENCE: 2
339 Phe Leu Ala Lys Glu Glu Ala Arg Lys Val Lys Arg Leu His Gly Met
340 1 5 10 15
343 Leu Arg Ser Leu Leu Val Tyr Met Leu Phe Leu Leu Val Thr Leu Leu
344 20 25 30
W--> 347 Ala Ser Tyr Gly Asp Ala Ser Cys His Gly His Ala Tyr Xaa Arg Leu
348 35 40 45
W--> 351 Gln Ser Xaa Xaa Xaa Xaa Ala Ile Lys Gln Glu Leu His Ser Arg
352 50 55 60
355 Ala Phe Leu Ala Ile Thr Arg Ser Glu Glu Leu Trp Pro Trp Met Ala
356 65 70 75 80
W--> 359 His Val Leu Leu Pro Tyr Val His Xaa Xaa Xaa Xaa Xaa Gly Asn Gln
360 85 90 95
W--> 363 Ser Ser Pro Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu
364 100 105 110
367 Gly Pro Pro Arg Leu Arg Gln Val Arg Leu Gln Glu Ala Leu Tyr Pro
368 115 120 125
371 Asp Pro Pro Gly Pro Arg Val His Thr Cys Ser Ala Ala Gly Gly Phe
372 130 135 140
375 Ser Thr Ser Asp Tyr Asp Val Gly Trp Glu Ser Pro His Asn Gly Ser
376 145 150 155 160
W--> 379 Gly Thr Trp Ala Thr Xaa Xaa Ser Ala Pro Asp Leu Leu Gly Ala Trp
380 165 170 175
383 Ser Trp Gly Ser Cys Ala Val Tyr Asp Ser Gly Gly Tyr Val Gln Glu
384 180 185 190
387 Leu Gly Leu Ser Leu Glu Glu Ser Arg Asp Arg Leu Arg Phe Leu Gln
388 195 200 205
391 Leu His Asn Trp Leu Asp Asn Arg Ser Arg Ala Val Phe Leu Glu Leu
392 210 215 220
395 Thr Arg Tyr Ser Pro Ala Val Gly Leu His Ala Ala Val Thr Leu Arg
396 225 230 235 240
399 Leu Glu Phe Pro Ala Ala Gly Arg Ala Leu Ala Ala Leu Ser Val Arg
400 245 250 255
403 Pro Phe Ala Leu Arg Arg Leu Ser Ala Gly Leu Ser Leu Pro Leu Leu
404 260 265 270
407 Thr Ser Val Cys Leu Leu Leu Phe Ala Val His Phe Ala Val Ala Glu
408 275 280 285
W--> 411 Ala Arg Thr Trp Xaa His Arg Glu Gly Arg Trp Arg Val Leu Arg Leu
412 290 295 300
415 Gly Ala Trp Ala Arg Trp Leu Leu Val Ala Leu Thr Ala Ala Thr Ala
416 305 310 315 320
W--> 419 Leu Val Arg Leu Ala Gln Leu Gly Ala Ala Asp Arg Gln Xaa Xaa Trp
420 325 330 335
423 Thr Arg Phe Val Arg Gly Arg Pro Arg Arg Phe Thr Ser Phe Asp Gln

```

RAW SEQUENCE LISTING ERROR SUMMARY  
 PATENT APPLICATION: US/09/753,008A

DATE: 12/14/2005  
 TIME: 12:15:53

Input Set : A:\96700-658.txt

Output Set: N:\CRF4\12142005\I753008A.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220>

to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 384,395,426,432,547,586,587,588,849  
 Seq#:2; Xaa Pos. 46,51,52,53,54,55,89,90,91,92,93,101,102,103,104,105,106  
 Seq#:2; Xaa Pos. 107,108,109,110,111,166,167,293,334,335,437,438,439,440  
 Seq#:2; Xaa Pos. 511,512,513,514,515,516,517  
 Seq#:3; Xaa Pos. 62,363,364,365,366,367  
 Seq#:4; Xaa Pos. 24,55,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,78,79  
 Seq#:4; Xaa Pos. 80,81,82,83,84,85,86,87,88,163,173,211,212,213,214,215,216  
 Seq#:4; Xaa Pos. 217,218,219,220,221,222,223,224,225,226,227,228,229,230  
 Seq#:4; Xaa Pos. 231,232,233,234,235,236,237,238,239,240,241,242,243,244  
 Seq#:4; Xaa Pos. 245,246,247,248,249,250,251,252,253,254,255,256,257,258  
 Seq#:4; Xaa Pos. 259,260,261,262,263,264,265,266,267,276  
 Seq#:6; N Pos. 3995,3996,3997,4906,4923,4932,4995,5028,5056  
 Seq#:13; N Pos. 11  
 Seq#:14; N Pos. 4  
 Seq#:15; N Pos. 10

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/753,008A

DATE: 12/14/2005

TIME: 12:15:53

Input Set : A:\96700-658.txt

Output Set: N:\CRF4\12142005\I753008A.raw

L:158 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:368  
L:162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:384  
L:170 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:416  
L:202 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:544  
L:210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:576  
L:278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:848  
L:347 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:32  
L:351 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:48  
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:80  
L:363 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:96  
L:379 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:160  
L:411 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:288  
L:419 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:320  
L:447 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:432  
L:463 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:496  
L:467 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:512  
L:501 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:48  
L:577 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:352  
L:641 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:16  
L:649 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:48  
L:653 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:64  
L:657 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:80  
L:677 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:160  
L:689 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:208  
L:693 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:224  
L:697 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:240  
L:701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:256  
L:705 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:272  
L:918 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:3960  
L:948 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:4860  
L:950 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:4920  
L:952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:4980  
L:954 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:5040  
L:1265 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0  
L:1280 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0  
L:1295 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0